WHAT IS CLAIMED IS:

1.   A method for making transmit/discard decisions on individual datagrams including the

steps of:

5   a) determining pipe membership of said datagrams by testing certain bits in said

datagram within each Input Interface Circuit of a network node upstream of the packet

forwarding device, the selection of bits made on the basis of standards for constructions

of datagrams with one or more of levels of headers and payloads;

          b) providing a table of values indicating transmit probability for each pipe of a

10   predetermined set within each said Input Interface Circuit;

          c) selecting a corresponding value matching said pipe membership of said datagram from

said table; and

          d) selectively transmitting or discarding said datagrams responsive to said transmit

probability matching said pipe membership.

15 2.   The method according to Claim 1, wherein the step of selectively transmitting further

includes the steps of:

e) providing a random number generator in the Input Interface Circuit;

f) performing a comparison of a current state of said random number generator with the value of transmit probability taken from said table of values;

g) discarding said datagram if the current state of said random number generator is closer than said value of transmit probability to a value representing a state where all datagrams of a pipe are to be transmitted; and

h) transmitting said datagram if the current state of said random number generator is not closer than said value of transmit probability to a value representing a state where all datagrams of a pipe are to be transmitted.

3. The method according to Claim 2, wherein said value representing a state where all datagrams of a pipe are to be transmitted is equal to one and said current state of said random number generator is a nonnegative fraction between 0 and 1.0.

4. The method according to Claim 1, wherein the step of pipe membership determination further includes the steps of comparing administratively specified values with the value of selected fields in a packet header of said datagrams.

5. The method according to Claim 4, wherein said selected fields of a packet header comprises the Differentiated Services Code Point (DSCP) field in a standard IP packet header.

6.   The method according to Claim 4, wherein said selected fields of a packet header

comprises the MPLS (Multiprotocol Label Switching) label or EXP bits or any combination

thereof, within the datagram, or ATM VPI/VCI fields also used for MPLS tunnel designation.

7.   The method according Claim 1 wherein said table of values is indexed according to pipe

5   numbers with each corresponding table entry representing the transmit probability corresponding

to the associated pipe.

8.   The method according to Claim 1, further comprising the step of altering said values

indicating transmit probability responsive to actual offered data rates in each of said pipes

relative to guaranteed minimum and / or maximum data rates of those pipes.

10 9.   The method according to Claim 8, wherein the step of altering said values is further

responsive to a signal from said packet forwarding device indicating congestion within that

device that affects one or more pipes flowing through said Input Interface Circuit.

10.   The method according to Claim 9, wherein said signal includes multiple components each

of which corresponds to congestion within a different output port or group of output ports, said

15 pipes within said Input Interface Circuit are grouped according to which output port or group of

output ports they direct data to, and wherein the step of altering said values is performed

separately for each group of pipes.

11. An apparatus in an Input Interface Circuit of a packet forwarding device comprising:

a) a table in which pipe numbers and associated probabilities of transmission are stored;

b) a buffer for storing at least a portion of a frame ;

c) a random number generator that periodically outputs random numbers;

5   d) a controller operatively coupled to the table, the buffer and random number generator, said controller:

parsing the information in the buffer to detect a pipe membership number for said frame,

determining from said table a transmission (transmit fraction) probability value

corresponding to the detected pipe membership number,

10   comparing the probability value with a present value of the random number generator and

discarding or forwarding the frame based upon the result of the comparison.

12. The apparatus according to Claim 11, wherein said controller alters said values indicating

transmit probability responsive to algorithm constants and actual offered data rates in each of

said pipes relative to guaranteed minimum and / or maximum data rates of those pipes.

13.   The apparatus according to Claim 12, further comprising a management component, said

management component initializing said algorithm constants and said guaranteed minimum and

maximum data rates of pipes.

14.   The apparatus according to Claim 11, further comprising an interface to an external

5   controller not contained within said Input Interface Circuit, said external controller being

connected to one or more Input Interface Circuit and altering said values indicating transmit

probability responsive to algorithm constants and actual offered data rates in each of said pipes

relative to guaranteed minimum and maximum data rates of those pipes within each of the Input

interface circuit it is connected to.

10   15.   The apparatus according to Claim 14, further comprising a management component, said

management component initializing said algorithm constants and said guaranteed minimum and

maximum data rates of pipes.

16.   The apparatus according to Claim 12, wherein said controller is further responsive to a

signal from said packet forwarding device indicating congestion within that device that affects

15   one or more pipes flowing through said Input Interface Circuit in altering said values.

17.   The apparatus according to Claim 14, wherein said external controller is further

responsive to a signal from said packet forwarding device indicating congestion within that

device that affects one or more pipes flowing through said Input Interface Circuit in altering said

values.

18. The apparatus according to Claims 16 or 17, wherein said signal includes multiple

components each of which corresponds to congestion within a different output port or group of

output ports, said pipes within said Input Interface Circuit being grouped according to which

output port or group of output ports they direct data to, and wherein the altering of said values is

5   performed separately for each group of pipes.

19. A method that controls the flow of datagrams including the steps of:

a) providing, in an Input Interface Circuit of a network node, a table identifying pipes

and associated transmission probability for each pipe in said table;

b) for each datagram received in said Input Interface Circuit determining the identity

10   of said datagram by examining bits within said datagram;

c) correlating datagram identity in step (b) with entries in the table; and

d) discarding or forwarding a datagram based upon the value of the transmission

probability only if the datagram identified in step (b) matches the identity of a pipe in the

table.

15 20. The method of Claim 19 wherein numerals are used to indicate the identity of pipes and

the identity of datagrams.

21.   The method of Claim 20 further including the steps of providing a random number

generated; and

discarding or forwarding a datagram based upon the value of the transmission probability

and a present value of the random number generator.

5   22.   A communication system comprising:

a switch;

a congestion monitoring device operatively coupled to said switch, said congestion

monitoring device monitoring transmission in said switch and generating at least one signal

indicating congestion state or no-congestion state in said switch;

10   a Physical Device (PHY) operatively coupled to the switch, said PHY receiving and

forwarding frames to said switch; and

a frame discard mechanism, located in said PHY, responsive to the at least one signal to

discard or forward frames to the switch based upon the state of the signal.

23.   A method to manage congestion in a communications network comprising the acts of:

15   operatively monitoring, with a congestion device, traffic within a communications

device;

generating at least one signal indicating congestion or no-congestion state in said

communications device; and

using the at least one signal to discard frames, prior to the communication device

initiating processing on the discarded frames, or forward frames to the communications device for processing.